

## GENOMICS ARTICLE

# Rose Scent: Genomics Approach to Discovering Novel Floral Fragrance–Related Genes<sup>W</sup>

Inna Guterman,<sup>a</sup> Moshe Shalit,<sup>a,b</sup> Naama Menda,<sup>a</sup> Dan Piestun,<sup>a</sup> Mery Dafny-Yelin,<sup>a</sup> Gil Shalev,<sup>a</sup> Einat Bar,<sup>b</sup> Olga Davydov,<sup>c</sup> Mariana Ovadis,<sup>a</sup> Michal Emanuel,<sup>a</sup> Jihong Wang,<sup>d</sup> Zach Adam,<sup>a</sup> Eran Pichersky,<sup>d</sup> Efraim Lewinsohn,<sup>b</sup> Dani Zamir,<sup>a</sup> Alexander Vainstein,<sup>a</sup> and David Weiss<sup>a,1</sup>

<sup>a</sup> Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agricultural, Food, and Environmental Quality Sciences, Hebrew University of Jerusalem, P.O. Box 12, Rehovot 76100, Israel

<sup>b</sup> Division of Aromatic Plants, Agricultural Research Organization, Newe Ya'ar, P.O. Box 1021, Ramat Yishay 30095, Israel

<sup>c</sup> Department of Plant Sciences, Weizmann Institute of Science, P.O. Box 26, Rehovot 76100, Israel

<sup>d</sup> Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109

For centuries, rose has been the most important crop in the floriculture industry; its economic importance also lies in the use of its petals as a source of natural fragrances. Here, we used genomics approaches to identify novel scent-related genes, using rose flowers from tetraploid scented and nonscented cultivars. An annotated petal EST database of ~2100 unique genes from both cultivars was created, and DNA chips were prepared and used for expression analyses of selected clones. Detailed chemical analysis of volatile composition in the two cultivars, together with the identification of secondary metabolism-related genes whose expression coincides with scent production, led to the discovery of several novel flower scent-related candidate genes. The function of some of these genes, including a germacrene D synthase, was biochemically determined using an *Escherichia coli* expression system. This work demonstrates the advantages of using the high-throughput approaches of genomics to detail traits of interest expressed in a cultivar-specific manner in nonmodel plants.

## INTRODUCTION

Flower fragrance is a composite character that is determined by a complex mixture of low-molecular-mass volatile molecules. For many years, research into flower fragrance focused on its chemical elucidation (Knudsen et al., 1993), and as a result, hundreds of compounds were identified. Most of these compounds belong to three major biosynthetic pathways: phenylpropanoids, fatty acid derivatives, and terpenoids (Croteau and Karp, 1991). Although the complete pathways leading to the final products have not been characterized, common modifications such as hydroxylation, acetylation, and methylation have been described (Dudareva, 2002).

No convenient plant model system that allows chemical and biochemical, as well as forward and reverse genetics, studies of flower scent is yet available. This is mostly the re-

sult of the invisibility of this character and its dynamic nature. Moreover, the well-established model system, *Arabidopsis*, generally is not suitable for this purpose, because the level of volatiles produced by its flowers is extremely low (Vainstein et al., 2001). As a result, the characterization of enzymes and genes involved in flower scent production and emission is still in its initial stages (Dudareva and Pichersky, 2000). To date, a modest number of floral scent genes have been isolated via classic biochemical approaches from plants with highly fragrant flowers, such as *Clarkia breweri* and *Antirrhinum majus* (Vainstein et al., 2001). Modern genomics, proteomics, and metabolomics tools can overcome the limitations inherent in classic studies on the mechanism of fragrance production and emission. Indeed, several groups recently used these high-throughput technologies to identify new fragrance genes in fruit and vegetative tissues (Aharoni et al., 2000; Lange et al., 2000; Gang et al., 2001).

Petals are the main source for fragrance compounds in many flowers, and fragrance production/emission is controlled developmentally. Based on expression analyses of fragrance-related genes, petal epidermal cells have been proposed as

<sup>1</sup> To whom correspondence should be addressed. E-mail weiss@agri.huji.ac.il; fax 972-8-9468263.

<sup>W</sup>Online version contains Web-only data.

Article, publication date, and citation information can be found at [www.plantcell.org/cgi/doi/10.1105/tpc.005207](http://www.plantcell.org/cgi/doi/10.1105/tpc.005207).

the sites for scent production and emission (Dudareva and Pichersky, 2000; Kolosova et al., 2001; Vainstein et al., 2001). Petal development in most plants can be divided into two phases: after initiation, petals exhibit a slow growth rate, resulting mainly from cell division; the later phase is characterized by rapid growth, resulting only from cell expansion (Martin and Gerats, 1993; Ben-Nissan and Weiss, 1996). Scent production and emission peak at the advanced stages of flower development, during cell expansion (Guterman et al., 2002). Other secondary metabolites involved in the attraction of pollinators, such as anthocyanins and carotenoids, also accumulate to high levels during this later phase (Vishnevetsky et al., 1999; Weiss, 2000).

For centuries, rose has been the most important crop in the floriculture industry. The genus *Rosa* includes 200 species and >18,000 cultivars (Weiss, 1997; Gudin, 2000). At an annual value of ~\$10 billion, roses are used as cut flowers, potted plants, and garden plants. Their economic importance also lies in the use of their petals as a source of natural fragrances and flavorings. The damask rose (*Rosa damascena*) is the most important species used to produce rose water, attar of rose, and essential oils in the perfumery industry. However, despite the importance of scent, most modern rose cut-flower varieties lack a distinct fragrance (Zucker et al., 1998). Because for many years selection for longevity was a major target and the loss of fragrance was observed, some negative correlation between these traits was suggested (Barletta, 1995). Nevertheless, the specific causes of the loss of fragrance during the course of breeding programs remain unknown. Several studies have been targeted at the volatile compounds emitted by old rose cultivars, and hundreds of different compounds representing the three main pathways—phenylpropanoids, fatty acid derivatives, and terpenoids—have been identified in different species and varieties (Flament et al., 1993). Thus, rose petals



**Figure 1.** Rose flowers of cultivars GG and FC at different developmental stages.

**Table 1.** Major Classes of Volatiles Emitted by FC and GG Flowers

Class of Volatiles	Amount ( $\mu\text{g}\cdot\text{flower}^{-1}\cdot 24\text{ h}$ )	
	FC	GG
Aromatic and aliphatic alcohols	$37 \pm 13$	(n.d. <sup>a</sup> )
Monoterpenes	$18 \pm 9$	(n.d.)
Sesquiterpenes	$10 \pm 3$	$0.3 \pm 0.2$
Esters	$61 \pm 42$	(n.d.)
Methoxylated phenols	$0.2 \pm 0.1$	$33 \pm 10$

Headspace was collected as described in Methods. Averages and standard errors of at least four replicates are given.  
<sup>a</sup>n.d., not detected.

represent an excellent system for the application of modern genomics technologies in the search for novel genes involved in fragrance production. Indeed, a recent genomics study aimed at identifying genes expressed in petals of *Rosa chinensis* has led to the putative identification of genes involved in scent formation (Channeliere et al., 2002).

Here, we describe the use of genomics to study rose petal genetic and biochemical functions. An annotated petal EST database was created from two *Rosa hybrida* tetraploid varieties with contrasting phenotypes, and expression profiling of selected clones was performed using microarray analyses. Detailed chemical analysis of volatile composition, together with the identification of secondary metabolism-related genes whose expression coincides with scent production, led to the discovery of several novel flower fragrance-related genes. Functional characterization of a few of these genes led to the identification of a sesquiterpene synthase involved in the production of germacrene D, two *O*-methyltransferases involved in the production of 3,5-dimethoxytoluene, and an alcohol acetyltransferase involved in the formation of geranyl acetate.

## RESULTS

### Rose Petal EST Database

A rose collection including scented “garden” varieties as well as commercial cut-flower cultivars was used to select cultivars for this study. The initial screening included a rough analysis of the volatile compounds produced by the flowers of the different cultivars, their petal pigments, and vase life (data not shown). Two cultivars with contrasting phenotypes were chosen for this work, Fragrant Cloud (FC) and Golden Gate (GG) (Figure 1). The large red FC flowers possess a strong scent, accumulate anthocyanins, and have a short vase life, whereas the small yellow flowers of GG accumulate carotenoids, have a long vase life, and lack a distinct

**Table 2.** Most Abundant ESTs in the Rose Database: The Largest Contigs Found and the Distribution of ESTs between FC and GG

Contig No.	Putative Function	ESTs in Contig	FC		GG	
			No. of ESTs	Percent of Total	No. of ESTs	Percent of Total
324	O-Methyltransferase	40	26	1.42	14	1.35
323	O-Methyltransferase	32	12	0.65	20	1.92
322	No homolog found	31	9	0.49	22	2.12
321	Unknown function	27	15	0.82	12	1.15
320	Low-molecular-mass heat-shock protein	19	12	0.65	7	0.67
319	Sesquiterpene synthase	17	17	0.93	0	0
318	Pollen allergen protein	10	5	0.27	5	0.48
317	$\delta$ -9 acyl-lipid desaturase	10	1	0.05	9	0.87
316	No homolog found	10	10	0.54	0	0
315	Metallothionein-like protein	10	5	0.27	5	0.48

odor (Guterman et al., 2002). Detailed gas chromatography-mass spectrometry (GC-MS) analyses of FC and GG flower headspace were performed to determine the type of volatiles emitted by these flowers (see supplemental data online), and the amounts of the major groups of volatiles are presented in Table 1. FC flowers emit numerous fragrance compounds known in roses (Flament et al., 1993), including aromatic alcohols, monoterpenes, sesquiterpenes, and various esters. Conversely, flowers of GG emit almost exclusively methoxylated phenols (99% of total volatiles). These compounds are hardly detected by the human nose. Analyses of the temporal regulation of volatile emission in FC and GG flowers revealed that the levels of most scent com-

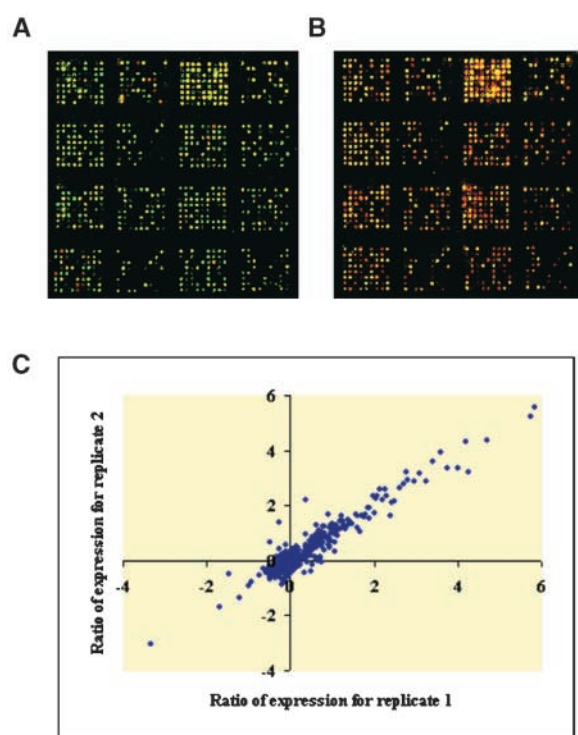
pounds peak at stages 4 to 6 (Figure 1) of flower development (data not shown).

To create the infrastructure to explore the relationships between sequences and phenotypes in rose, we generated two directionally cloned cDNA libraries from FC and GG stage-4 petals. Randomly chosen individual clones were sequenced from the 5' terminus, and vector sequences, low-quality 3' sequences, and artifacts were edited out. High-quality sequence information of 2873 individual clones, 1834 from FC and 1039 from GG, was annotated and organized in contigs (see supplemental data online). A very high number of unique genes (2139; 74% of total ESTs) is represented in the database. Among these, 1288 genes were

**Table 3.** Functional Classification of Rose Petal Genes

Functional Group	No. of Genes (FC + GG)	Percent of Total (FC + GG)	Percent of FC Database	Percent of GG Database
Unclassified	681	32	32	31
Homolog not found	365	17	17	17
Metabolism	329	15	15.8	14.6
Signal transduction	134	6	6.4	5.7
Protein destination	92	4.3	4.6	3.5
Cell rescue/defense	87	4	4	4.3
Protein synthesis	83	4	3	5.5
Energy	64	3	2.8	3.5
Transport facilitation	55	2.6	2.9	2.2
Transcription	54	2.6	2.5	2.6
Cellular organization	51	2.5	2.1	3.1
Cell growth/division	37	1.7	1.6	1.9
Cellular biogenesis	31	1.5	1.4	1.6
Development	28	1.4	1.2	1.8
Intracellular transport	24	1.2	1.3	0.9
Ionic homeostasis	14	0.8	0.6	0.7
Motility	5	0.2	0.4	0.1
Plasmid proteins	5	0.2	0.4	0

Unique genes (from FC and GG) were classified into functional groups according to the Munich Information Center for Protein Sequences.



**Figure 2.** Evaluation of the Microarray Analysis.

**(A)** Overlay image of the microarray hybridized with labeled cDNA originating from FC petals at stage 4 (Cy3) and GG petals at the same stage (Cy5).

**(B)** Overlay image of the same microarray as in **(A)** hybridized with the same probes but with reverse labeling.

**(C)** Scatterplot of average signal values of three replicates on the microarray. Poly(A)<sup>+</sup> RNA was isolated from FC petals at stage 4 and from GG petals at the same stage, reverse transcribed, and labeled with Cy5 and Cy3. In the first experiment, FC RNA was labeled with Cy3 and GG RNA was labeled with Cy5, and in the second experiment, labeling was reversed. The  $\ln$  values of the Cy5-to-Cy3 ratios were plotted.

found only in the FC database and 746 were found only in the GG database. Most unique genes (1815; ~63% of total ESTs) were found as singletons, indicating low redundancy of libraries and suggesting that rose petals are a very rich source for gene mining. Similar redundancies were found in FC- and GG-specific databases. Thirty-seven percent of the ESTs belong to contigs, ranging from low redundancy (2 to 5 ESTs per contig, 299 contigs detected) to medium redundancy (6 to 9 ESTs per contig, 15 contigs found) to high redundancy ( $\geq 10$  ESTs per contig, 10 contigs found). Although most of the highly redundant contigs contain ESTs from both varieties, two contain ESTs from only FC (Table 2), suggesting cultivar-specific expression. The largest contig in the database, showing sequence similarity to O-meth-

yltransferase (OMT), contains 40 ESTs with equal representation for both cultivars.

The unique genes were classified into 18 functional groups (Table 3) according to the Munich Information Center for Protein Sequences (<http://mips.gsf.de>). Separate analysis of the databases for the two cultivars revealed that the distribution of unique genes among functional groups is essentially identical in the FC and GG databases. The largest group of petal sequences, containing approximately one-third of all unique genes (32%), show homology (Expect value  $< 1.0 \times 10^{-5}$ ) with genes coding for predicted proteins with unknown function. The second largest group (17%) contains unique sequences that have no homologs (Expect value  $\leq 1.0 \times 10^{-5}$ ), as revealed by BLAST analysis against NR (the nonredundant proteins database), all plant ESTs (EST\_pln database), and the Arabidopsis genomic sequences. The high percentage of previously unknown sequences suggests that rose petals are an intriguing source for novel genes. Another large group of genes in the rose database (15%) is represented by proteins putatively involved in metabolism. Among these, 233 unique genes are related to primary metabolism and 96 (~5% of total ESTs) code for enzymes putatively involved in secondary metabolism.

### Expression Analyses of Rose Petal Genes

To identify genes whose expression coincides with scent production, microarray analyses were performed with 350 selected unique genes from FC. ESTs printed on slides were of the following putative functions: primary and secondary metabolism, development, transcription, cell growth, cell biogenesis and organization, cell rescue, signal transduction, and ESTs with unknown function. Each glass slide held three copies of the entire array. To ensure the reliability of the results, two microarray slides (six replicates) were used for each experiment. Two independent RNA preparations were made for each analysis, and labeling of the cDNA (Cy3 versus Cy5) was reversed on the second slide. RNA was extracted from FC and GG petals at stage 4 and used as probes to compare gene expression between the two cultivars. Figures 2A and 2B show the two arrays (FC stage 4 versus GG stage 4) with reciprocal labeling, and the scatterplot of signals from these two arrays is presented in Figure 2C. The correlation coefficient of the ratios from these two arrays was 0.95, suggesting good reproducibility among individual arrays in the same experiment. Using a twofold change in gene expression as the threshold (Ma et al., 2001), 77 genes (22% of printed cDNAs) showed higher expression in FC petals than in GG petals (Figure 3A). Approximately 270 genes (~77%) showed similar expression in the two varieties, and 3 genes (1.0%) showed higher expression in GG petals. In a second set of experiments, the expression level of the selected genes on the chips was compared between FC petals at stages 1 and 4. These results show that 65 genes (18.5%) were upregulated during FC petal devel-

opment and 14 genes (4%) were repressed (Figure 3B). Approximately 270 genes (~77%) exhibited similar expression levels in young (stage 1) and developed (stage 4) FC petals.

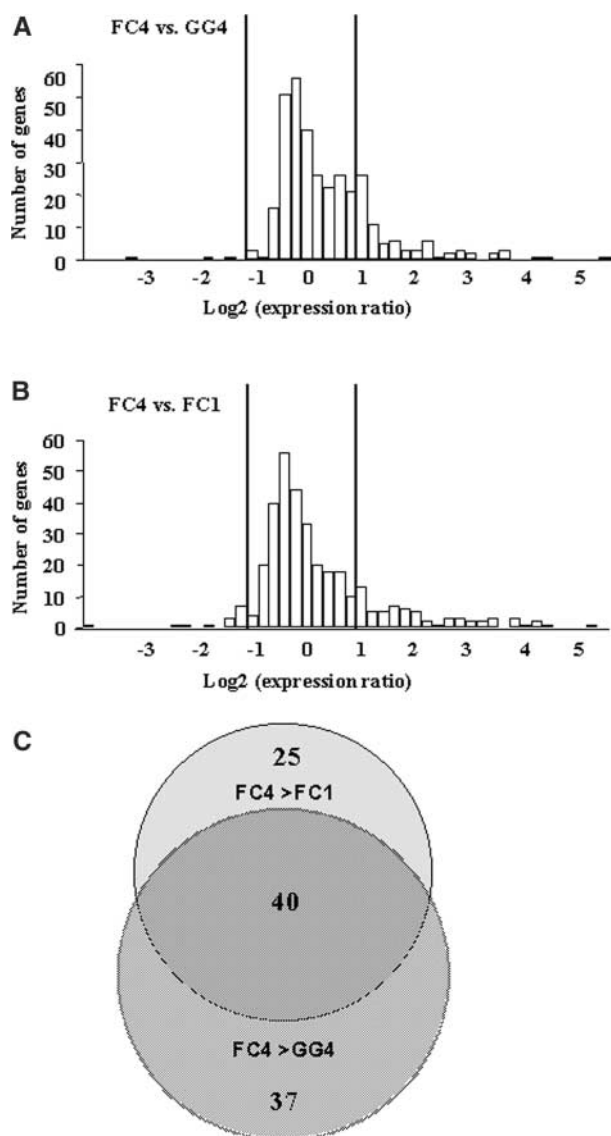
Because most scent compounds are produced by FC petals at stage 4 but not at stage 1 (Guterman et al., 2002) and not by GG petals, we overlapped ESTs whose expression is upregulated during FC petal development with those that show higher expression in stage-4 FC versus stage-4 GG petals (Figure 3C). This analysis yielded 40 unique genes that were upregulated during FC petal maturation and exhibited higher expression in FC than in GG flowers (see supplemental data online). Fifteen of these genes (38%) were classified into the metabolism functional group, seven with a putative role in secondary metabolism (Table 4). Although five of them showed relatively moderate differential expression, FC0284 and FC0592 exhibited very strong up-regulation. These clones are similar to the previously identified glutamate decarboxylase and (+)- $\delta$ -cadinene synthase (a sesquiterpene synthase) genes, respectively.

Because FC and GG petals differ in size and color, it was not surprising that genes putatively involved in these traits also were represented among the 40 clones found in the overlapping analysis (Table 4). For example, genes related to cell growth included a putative xyloglucan endotransglycosylase (Schünmann et al., 1997) and a GAST-like gene (Shi et al., 1992). Additionally, genes possibly involved in anthocyanin production included a putative dihydroflavonol reductase and a gene similar to the transcriptional regulator myb26 (Uimari and Strommer, 1997).

Because FC petals produce and emit high levels of sesquiterpenes (Table 1), we characterized in detail clone FC0592, whose sequence shows similarity to genes of the sesquiterpene synthase family (Figure 4). This clone is 2092 bp long and bears an open reading frame coding for 565 amino acids. The deduced protein sequence exhibited the highest similarity (53% identity) to the deduced sequence of the (+)- $\delta$ -cadinene synthase from cotton, a sesquiterpene synthase involved in the biosynthesis of antifungal compounds (Chen et al., 1995). FC0592 belongs to one of the largest contigs in the database, containing 17 ESTs, all from FC (Table 2), suggesting cultivar-specific expression. To further characterize the spatial and temporal expression of this gene, RNA gel blot analysis was performed (Figure 5). A high level of expression was detected in mature FC petals (stages 4 to 6) but not in young petals (stage 1) or in young or mature leaves. In GG, no expression of this gene was detected in any of the tissues analyzed, including mature petals.

### Functional Analyses of Rose Scent-Related Genes

To determine whether this clone codes for a sesquiterpene synthase and to identify the type of sesquiterpene produced, the entire reading frame of FC0592 was cloned



**Figure 3.** Expression Profiles of Selected Rose Genes.

(A) Distribution of rose genes according to the ratio of expression in stage-4 FC versus stage-4 GG petals. The ratio for each clone is the average of six replicates taken from two independent experiments. Threshold values are indicated by vertical lines.

(B) As in (A), but comparing expression in FC petals at stage 1 versus stage 4.

(C) Interlocking diagram of differentially expressed genes that exhibited increased expression during FC petal maturation and higher expression in mature (stage 4) FC versus GG petals. The threshold for increased expression was twofold. The number in the overlapping area indicates the number of unique genes that exhibited upregulation in both experiments.

**Table 4.** A Partial List of Genes Whose Expression Is Upregulated during FC Petal Maturation and Is Higher in Mature FC Petals Than in Mature GG Petals, as Revealed by Microarray Analyses

Clone No.	Accession No.	Putative Function	Expression (Fold Increase)	
			FC4/FC1	FC4/GG4
		Secondary metabolism		
FC0284	BQ104391	Glutamate decarboxylase	22.2	45.3
FC0445	BQ105103	Methylglutaryl-CoA reductase	9.5	3.3
FC0524	BQ105675	Dihydroflavonol reductase	2.8	2.1
FC0592	BQ106271	Sesquiterpene synthase	14.4	23.4
FC1173	BQ105028	Cytochrome P450-like	2.1	2.5
FC1386	BQ105113	Tyr decarboxylase	2.3	2.7
FC1447	BQ104118	Cytochrome P450-like	3.4	2.1
		Transcription		
FC0293	BQ106609	GhMyb9-like	2.1	2.0
FC0568	BQ106505	Myb26-like	14.8	2.0
		Cell growth		
FC0620	BQ104195	Xyloglucan endotransglycosylase	16.3	2.5
FC0861	BQ105983	GAST-like	5.2	9.3
FC1035	BQ104371	Aquaporin-like	20.2	2.0

into an expression vector to obtain a catalytically active recombinant enzyme. In vitro assays using lysates from bacteria overproducing the FC0592 gene product and the substrate farnesyl diphosphate (FPP) resulted in the synthesis of a sesquiterpene, identified by GC-MS as germacrene D (Figure 6). Germacrene D was not detected in control assays prepared from lysates derived from bacterial cells expressing an unrelated gene. Figure 6C shows the structures of the substrate FPP and the product germacrene D. Under our assay conditions, no monoterpenes were detected when geranyl diphosphate was presented as a substrate to the FC0592 gene product. However, the ability of this protein to use geranyl diphosphate in vivo cannot be excluded.

Headspace analysis of FC and GG petals revealed the emission of sesquiterpenes in both varieties (Table 5). Interestingly, the few sesquiterpenes identified in the headspace of GG were not detected in the headspace of FC flowers and vice versa. Nevertheless, the overall level of sesquiterpenes emitted by FC flowers was much higher (~30 times) than that emitted by GG flowers. Similarly, sesquiterpene synthase activity in cell-free extracts was much higher in FC than in GG petals (6.0 versus 0.2 fkat/mg protein, respectively; data not shown). The major sesquiterpene identified in the headspace of FC flowers was germacrene D (~60% of the emitted sesquiterpenes). The emission of germacrene D during flower development was evaluated by headspace analysis. Germacrene D emission from FC petals increased sharply during flower opening, reaching a maximum value at stage 4 (Figure 7). By contrast, no germacrene D was detected in the headspace of GG flowers at any flowering stage.

To search for additional scent-related genes, we looked for ESTs whose expression increases during FC petal devel-

opment, as revealed by the microarray analysis, regardless of their expression level in GG. Among these 88 unique genes, those with putative functions in secondary metabolism included a putative alcohol acetyltransferase, two putative OMTs (OMT1 and OMT2), and sequences with strong similarity to known monoterpene synthases, decarboxylases, hydroxylases, aminotransferases, and aldehyde dehydrogenases (Figure 8). RNA gel blot analyses of selected clones confirmed the results of the microarray: their expression level increased during FC petal maturation. For some ESTs, upregulation from stage 1 to stage 4 also was found in GG petals (Figure 8). The functions of the putative alcohol acetyltransferase and the two OMTs were confirmed in vitro, using lysates from *Escherichia coli* expressing these genes. The putative alcohol acetyltransferase catalyzed the transfer of the acyl group of acetyl-CoA to several alcoholic acceptor substrates, having the highest catalytic efficiency toward geraniol-producing geranyl acetate (our unpublished data). Both OMTs could methylate a variety of phenolic substrates using S-adenosylmethionine as the methyl donor. OMT1 exhibited the highest catalytic efficiency toward orcinol and produced orcinol methyl ether, and OMT2 used orcinol methyl ether most efficiently to produce orcinol dimethyl ether (Lavid et al., 2002). It should be noted that geranyl acetate and orcinol methyl and dimethyl ether are found in the headspace of FC and GG flowers (see supplemental data online).

## DISCUSSION

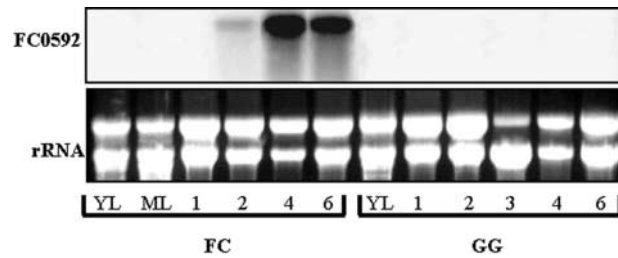
In the last decade, two plant systems, *Clarkia breweri* (Dudareva and Pichersky, 2000) and *Antirrhinum majus*

(Dudareva et al., 2000), have been used extensively to identify and characterize floral scent genes. Because forward genetics approaches could not be harnessed (Vainstein et al., 2001), these studies have used biochemical tools (i.e., from enzyme activity through protein purification to gene cloning; Dudareva et al., 1996, 1998). More recently, genomics approaches have been adapted to study genes that control the biosynthesis of both floral and vegetative volatiles, because these new methods allow quick access to plants with poor genetic characterization (Aharoni et al., 2000;

Rose	M S L N V V S V A S Q A P P T L N A S T D A G S R S S A N F H P S I V G D O F L	40
Cotton	. . . . M A S Q A S Q V L A S P H P A I S S E N R P K A D F H P C I V G D H F I	36
Tomato	. . . . . M A A S S A D K C D P L A N F H P S V W G Y H F L	25
Rose	S Y N N T M E I N I N G E Q H V Q Q L K E E V K R M L H A P P V Q T S L G K L E	80
Cotton	I C P D I D I A A T E L Q Y E L K A Q V R K H I M E . . P V D D S N Q K L P	74
Tomato	S Y T H E I T N Q K V E V D E Y . . K E T I R K M L V E T C D N S T Q . K L V	62
Rose	L I D D I Q R L C V S Y H F E Y E I D Q T L Q Q T H N F N G Y C D D . D L H T	119
Cotton	F I D A V Q R L C V S Y H F E K E T E D E L E N T Y R D T N N N D A D T D L Y T	114
Tomato	L I D A M Q R L G V Y H F D N E I E T S I Q N T F D A S S K Q N D N N L Y	102
Rose	C A L R F E L L R O H C Y N V S C D T E N R F K D S D C K L N K S L H D V V	158
Cotton	T A L R F E L L R E H C F D I S D D A E N K F K D E A C N F K A S L T S D V Q	153
Tomato	V V S L R F R L V R Q Q C Y M S S D V E R Q E T N Q D C K F K E T L E N D V Q	142
Rose	G L Q S L Y E A T H L R V R G E D F L E E A L T T T T H L A S A V L A N R L S	198
Cotton	G L E L Y E A S Y M R V H C G D I L D E A I S Y T T A Q L T L A L P T L H H .	192
Tomato	G L S L Y E A S H L R V R N E L L E E A L T T T T H L E S I V S N L S N N	182
Rose	P P L S K Q V R H . A L N Q D L R C G L P L E A R H Y M S L H Q E L H G S R N	237
Cotton	. P L S E Q V C H . A L K Q S I R R C L P R V E A R N F I S Y Q D L S H N K	230
Tomato	N N S L R V E V C E A L T Q D I D M T L P M C A R K Y I S I Y E N D A H H H	222
Rose	Q I L L S F A K L D F N L L Q Q L H Q K E L S D I A R V U K E L D F A S K D F F	277
Cotton	S . L L Q F A K I D F N L L Q L H R K E L S E I C R V U K D L D F T R K D F F	269
Tomato	L . L L K F A K L D F N L L Q K F H Q R E L S D L T R W M K D L D F A N K Y F F	261
Rose	A R D R E I C Y F W I L C V Y F E P K Y F A R T I L T K V E A M T S I I D D	317
Cotton	A R D R V V C Y F W I M C V Y F E P Y S L C R M M T K V E A M A S I V D D	309
Tomato	A R D L V E C Y F W I L C V Y F E P R Y S R A R M M T K V E N L T S I I D D	301
Rose	I Y D V Y C L E E D L F T R A T E R W D I S A M D L P E Y M K V C Y R A L	357
Cotton	T Y D S Y A T D E L I P Y T N A T E R W D I K C H N Q L P N Y M K I S Y K A L	349
Tomato	T F D A Y A T F D E L V T E N D A T O R W D A L A I D S I Q Y M R P A Y A L	341
Rose	L D V Y A D T E Q L A S K C R L Y B I D Y A K E A M K R O V T A Y H A E A K U	397
Cotton	L D V Y E E H E Q L A N Q C R O Y P R V Y A K A M I P L V D A Y L L E A K U	389
Tomato	L D I Y S E H E Q V L S K E C K L D R V Y A K E M K L V A Y F K E T Q U	381
Rose	F H K N Y T P T M D R Y H E V A L V T S A Y S H L A . T T S F V C M G D I V T K	436
Cotton	T H Q N Y K P T F E E F R D N A L P T S C Y A M L A . I T A F V C M G E V I T P	428
Tomato	L N D C D H I P K Y E R Q V E N A I V S A G Y M H I S T T C L V G I E F I S H	421
Rose	D S F E M I F S D P H I V K A S S V V C F E M D D I V S H K F E Q R E G H V A S	476
Cotton	E T F E V A A S D P H I K A S T I I C F E M D D I A E H K E N H R E D D C S	468
Tomato	E T F E V L N H S V L V R A S A L I A E A M D I V G H E D E Q R E G H V A S	461
Rose	A V E C Y M K Q Y C A T E E Q T I I E F P R Q V S D A U K D E N E E C L H P T P	516
Cotton	A I E C Y M K Q Y C T A Q E A Y N E F N K H I E S S W K D Y N E E F L K P T E	508
Tomato	L I E C Y M K Q Y G A S K Q E T I Y K F F E E V T N A U K D I N K Q F S R P T E	501
Rose	L D H P L L I R V L N L T R E V I D W L Y K H E D Y T H A G T I L K D F V A S L	556
Cotton	M P T P V L R S L N L A R V H D V L Y E C D C Y T H V G K A A K C G T I S L	548
Tomato	V P H F V L R V L N L T R A V A D L Y K E K D Y T S A K C K L R H M I N P I	541
Rose	L I E P V P V P V	565
Cotton	L I D F I Q I	555
Tomato	L I E S V K I	548

**Figure 4.** Alignment of the Deduced Amino Acid Sequence of FC0592 (the Putative Rose Germacrene D Synthase) with Those of Cotton (+)- $\delta$ -Cadinene Synthase and Tomato Germacrene C Synthase.

Amino acids conserved in all three sequences are shaded black, and those conserved in only two sequences are shaded gray.



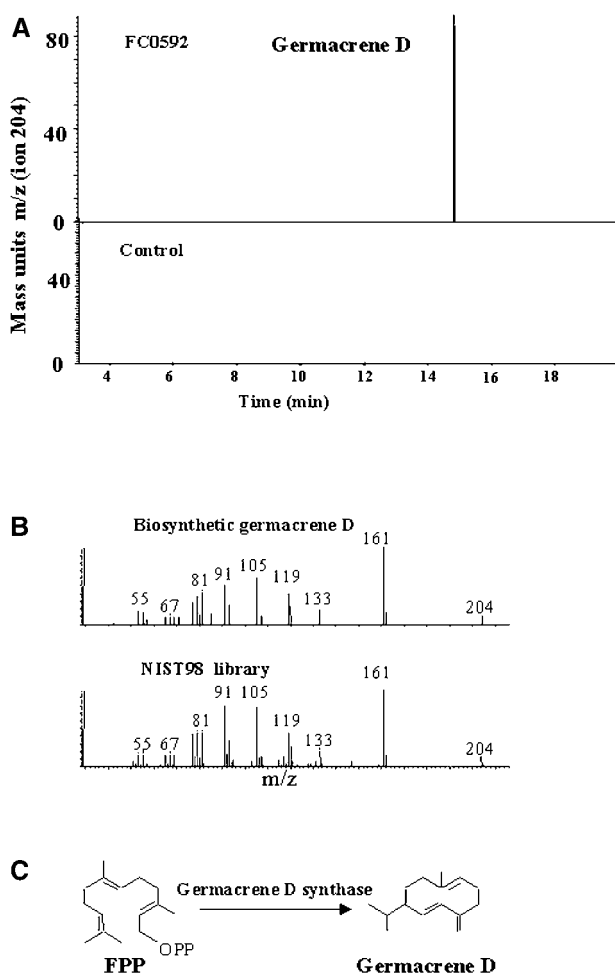
**Figure 5.** RNA Gel Blot Analysis of the Putative Sesquiterpene Synthase.

Total RNA was extracted from FC and GG young (YL) and mature (ML) leaves and from FC and GG petals at different developmental stages (1, 2, 4, and 6) and analyzed for FC0592 expression.

Lange et al., 2000; Gang et al., 2001; Channeliere et al., 2002). Moreover, these modern tools enable one to choose a plant system to study according to the trait of interest rather than being limited to established model systems (Fiehn et al., 2000). Here, we used the tools of genomics to identify novel scent genes in rose petals. By combining EST database mining with metabolic profiling and microarray expression analyses, we were able to identify several novel genes involved in fragrance production in rose petals.

### Rose Petal ESTs

The petals, with their pigments and scents, are the major contributors to the attractiveness of rose flowers; as such, they are responsible for the economic value of the plant. We created an annotated petal EST database from two *R. hybrida* cultivars with contrasting phenotypes: an old garden variety with large red fragrant flowers that have a short vase life (FC), and a modern cut-flower cultivar with small yellow flowers and only faint odor but a long vase life (GG). The petal cDNA libraries used in this study for the generation of the EST database were not normalized, allowing an estimation of the expression level of specific genes based on the corresponding contig sizes. Furthermore, the existence of a specific contig in only one variety suggested cultivar-specific expression. Indeed, for several analyzed clones, we found a good correlation between the number of ESTs in a contig and the level of their transcripts. For example, the sesquiterpene synthase contig is one of the largest in the database, containing 17 ESTs, all from FC. High expression of this gene in FC but not in GG petals was confirmed by RNA gel blot analysis. In addition, two OMTs were found in two large contigs. These contigs contained ESTs from both FC and GG: 12 ESTs from GG and 20 ESTs from FC represented OMT1, and 14 GG ESTs and 26 FC ESTs represented OMT2. RNA gel blot analysis confirmed the high expression of these OMTs in both cultivars (Figure 8).



**Figure 6.** Functional Analysis of Clone FC0592 in *E. coli*.

Lysates extracted from *E. coli* overexpressing clone FC0592 were incubated with FPP in assay buffer, and the reaction product was assayed by GC-MS.

**(A)** Single-ion GC-MS chromatogram of hexane extracts of bacterial lysates derived from cells overexpressing FC0592 or from cells overexpressing FC1018 used as a control.

**(B)** The sesquiterpene produced had a mass spectrum corresponding to that of germacrene D (99% identity) compared with the computerized NIST98 library. The Kovac index of this compound also corresponded with that of authentic germacrene D. These properties were identical with those of authentic germacrene D present in the commercial essential oil of *Cananga odorata*.

**(C)** Chemical structures of FPP and the reaction product germacrene D.

Database annotation for both cultivars revealed relatively low redundancy (Sterky et al., 1998), with more than ~60% of the total ESTs found as singletons. This suggests a large collection of genes with similar representation in the library and, therefore, possibly similar expression in the tissue. The

**Table 5.** Sesquiterpenes Emitted from FC and GG Flowers

Sesquiterpene	Amount ( $\mu\text{g}\cdot\text{flower}^{-1}\cdot 24\text{ h}$ )	
	FC	GG
$\alpha$ -Cubebene	$1.2 \pm 1.1$	(n.d. <sup>a</sup> )
$\beta$ -Elemene	$0.1 \pm 0.1$	(n.d.)
<i>trans</i> -Caryophyllene	(n.d.)	$0.3 \pm 0.1$
Germacrene D	$6.2 \pm 1.6$	(n.d.)
Bicyclosquisphellandrene	$0.2 \pm 0.1$	(n.d.)
$\alpha$ -Murolene	$0.2 \pm 0.1$	(n.d.)
$\delta$ -Cadinene	$2.1 \pm 1.5$	(n.d.)
$\alpha$ -Copaene	$0.2 \pm 0.1$	(n.d.)
$\beta$ -Cubebene	$0.05 \pm 0.01$	(n.d.)
Bicycloelemene	$0.05 \pm 0.01$	(n.d.)
Bicyclogermacrene	$0.1 \pm 0.05$	(n.d.)
Caryophyllene oxide	(n.d.)	$0.05 \pm 0.01$
$\alpha$ -Farnesene	(n.d.)	$0.05 \pm 0.01$
Total sesquiterpenes	(10.4)	(0.4)

Headspace was collected as described in Methods. Averages and standard errors of at least four replicates are given.

<sup>a</sup>n.d., not detected.

expression of numerous unique genes in concert may be required for both rapid growth of the petals and the production of many compounds from primary and secondary metabolism. Indeed, cataloging the unique genes into functional groups revealed a large collection of proteins that are putatively involved in metabolism (~16%) and cell growth and organization (~5%). More than 30% of the proteins in the first group are likely to be involved in secondary metabolism. Moreover, genes related to secondary metabolism seem to be expressed at very high levels in the mature petals, as indicated by contig size: almost half of the largest contigs ( $\geq 10$  ESTs) in the rose database contain genes putatively involved in secondary metabolism. It also is likely that additional cDNAs involved in secondary metabolism are found among the group of novel unknown sequences, because genes for secondary metabolism are likely to be the most divergent among plant genomes and therefore more difficult to identify as such by sequence comparisons alone (Pichersky and Gang, 2000).

Many EST databases using model as well as economically important plant systems have been generated in the last several years. Although various plant organs were used for these analyses, information on ESTs specific to petals has just begun to accumulate (Vainstein et al., 2001; Channeliere et al., 2002). The limited number of published petal-specific ESTs may explain the relatively large number of rose petal ESTs (~17%) representing novel putative proteins with no significant similarity ( $E$  value  $\geq 10^{-5}$ ) to sequences in the database. The observation that some of these ESTs were found in contigs containing both FC and GG sequences supports the authenticity of these proteins. Similar results were found recently by Channeliere et al. (2002), who used petals of *R. chinensis*, an ancestral spe-

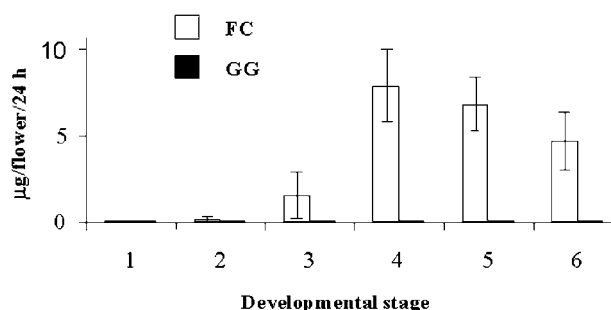
cies of *R. hybrida*, to generate an EST database. A large group of novel sequences also were found in the petal EST databases of *C. breweri* (13%) (<https://sativa.biology.lsa.umich.edu/blast/blast.html>; login, clarkia; password, brewery). Comparison between these novel sequences of *C. breweri* and rose revealed only one match. The rapid progress in the accumulation of genomic sequences from various species, as well as that of ESTs from various organs/specialized tissues, should reveal whether these novel putative proteins are petal-specific and/or unique to roses. To this end, the accessibility of the *R. chinensis* database for data mining should be highly instrumental.

Although petals of FC and GG are phenotypically different, the distribution of ESTs isolated from the two cultivars between the various functional groups is almost identical. Thus, FC and GG petals probably express a similar set of genes that determine the nature of the organ. Their final appearance, including scent, on the other hand, may be determined by a relatively small number of genes that are expressed in a cultivar-specific manner. Based on expression profiling and sequence annotation, several genes with putative functions in secondary metabolism in general and scent production in particular were identified (Table 4, Figures 6 and 8). Among these are genes putatively involved in modification reactions such as hydroxylation, acetylation, and methylation. Using an *E. coli* expression system, the biochemical functions of alcohol acetyltransferase (our unpublished data), two orcinol methyltransferases (Lavid et al., 2002), and sesquiterpene synthase (Figure 8) were confirmed.

### Rose Scent-Related Genes

One of the main groups of volatiles emitted by FC flowers is the sesquiterpenes. Similarly, one of the largest contigs in the database contains a sesquiterpene synthase gene. No homologous sequences were found in the GG database. Microarray analysis, further confirmed by RNA gel blot analysis, revealed that this gene is upregulated strongly during FC petal development and is not expressed in GG petals. The encoded sesquiterpene synthase protein produced the sesquiterpene germacrene D in vitro. This volatile is the most abundant sesquiterpene produced by FC petals. Germacrenes are common constituents of the aromas of many plants, and they are regarded as bound or free intermediates in the biosynthesis of other sesquiterpenes (Colby et al., 1998).

Sesquiterpenes are the most diverse class of terpenoids, and many important sesquiterpenes are formed by plants as flavor and aroma constituents (e.g., farnesol, nerolidol, caryophyllene, and nootkatone) (Croteau and Karp, 1991). Numerous sesquiterpene synthases, including the tomato germacrene C synthase, have been isolated in the last few years (Colby et al., 1998; Bohlman et al., 2000). Sesquiterpenes also serve as phytoalexins, and genes for several sesquiterpene synthases involved in phytoalexin biosynthesis have been isolated and characterized, including the to-



**Figure 7.** Germacrene D Emission from Rose Petals during Development.

Germacrene D levels were measured using the headspace technique described in Methods. FC and GG flowers were sampled at different stages of flower development. Each point represents an average of three to five individually sampled flowers. Germacrene D was identified by GC-MS.

bacco 5-epi-aristolochene synthase (Facchini and Chappell, 1992), the *Hyoscyamus muticus* vetispiradiene synthase (Back and Chappell, 1995), and the cotton (+)- $\delta$ -cadinene synthase (Chen et al., 1995).

Sesquiterpene synthases produce diverse sesquiterpenes from the ubiquitous precursor FPP. Sequence analyses have revealed that although conservation is not high among terpene synthase genes from different plant species, discrete conserved domains are present, suggesting significant structural and functional similarity (Back and Chappell, 1995; Bohlman et al., 2000). Moreover, similarity between genes within this group does not necessarily determine similarity between their products (Bohlman et al., 2000). Indeed, the molecular structure of germacrene D (Figure 6C) resembles that of its isomer germacrene C (Colby et al., 1998). However, rose germacrene D synthase (FC0592) shows greater similarity to cotton (+)- $\delta$ -cadinene synthase, which catalyzes the synthesis of a sesquiterpene with a different skeleton, than to tomato germacrene C synthase.

### Linking Rose Sequences with Phenotypes

An intriguing goal for a future study is the identification of members of the complex machinery that regulates scent production, including the elucidation of network links between pathways leading to the production of different groups of volatiles. To date, such elements have not been identified, mainly because forward genetics cannot be used because of the lack of an efficient screening method for scent mutants. The use of genomics tools may allow the identification of such elements. For example, analysis of our microarray revealed two Myb-like genes whose expression is upregulated during FC petal maturation and is higher in mature FC petals than in GG petals. One of these genes,

Clone number	Accession number	Putative function/product	Developmental stage (RNA)			
			FC		GG	
			1	4	1	4
FC1018	BQ106456	Alcohol acetyl transferase/geranyl acetate <sup>a</sup>				
FC0606	BQ105775	Alcohol dehydrogenase			n.a.	
FC0109	BQ105581	Aldehyde dehydrogenase	n.a.		n.a.	
FC1082	BQ105569	Aspartate transaminase			n.a.	
FC1081	BQ104799	Cytochrome P450	n.a.		n.a.	
FC0284	BQ104391	Glutamate decarboxylase				
FC1140	BQ105123	Glutamate decarboxylase	n.a.		n.a.	
FC0496	BQ106549	O-methyltransferase-1/ orcinol methylether <sup>b</sup>				
FC0895	BQ106672	O-methyltransferase-2/ orcinol dimethylether <sup>c</sup>				
FC0444	BQ104299	Tyrosine decarboxylase-1	n.a.		n.a.	
FC1386	BQ105113	Tyrosine decarboxylase-2			n.a.	
FC1079	BQ105216	Monoterpene synthase				

**Figure 8.** Rose Genes with a Putative Role in Secondary Metabolism Whose Expression Is Upregulated during FC Petal Maturation, as Revealed by Microarray and/or RNA Gel Blot Analysis.

n.a., not analyzed. <sup>a</sup> Gene functionally identified as an alcohol acetyl-CoA:acetyltransferase. <sup>b</sup> Gene functionally identified as an S-adenosylmethionine:orcinol methyltransferase (OMT1). <sup>c</sup> Gene functionally identified as an S-adenosylmethionine:orcinol methyl ether methyltransferase (OMT2).

FC0568, shows the greatest similarity to pea Myb26 and to snapdragon Myb305 and Myb340. These Myb proteins regulate phenylpropanoid biosynthesis and have a high affinity for promoters of phenylpropanoid genes (Uimari and Strommer, 1997). The rose homolog also may regulate phenylpropanoid biosynthesis in petals, leading to the production of flavonoid pigments, as was found in snapdragon (Sablowski et al., 1994). However, activation of the phenylpropanoid pathway also may lead to the enhanced production of scent compounds originating from this pathway, such as phenyl ethyl alcohol and benzyl alcohol, which are produced by FC petals (see supplemental data online). Cross-talk between the branches of the phenylpropanoid pathways leading to pigment and scent production has been demonstrated recently (Zuker et al., 2002).

The cultivated rose is a unique crop plant that has benefited from thousands of years of breeding and human selection. In contrast to many other cultivated plants that have a narrow genetic basis as a result of a strong founder effect during domestication (Zamir, 2001), the *R. hybrida* genome includes footprints for introgressions from at least 10 wild

species (Gudin, 2000). This wealth of phenotypic variation, combined with the tools of genomics, can lead to the association of sequences and phenotypes, as was shown in this study for the varieties GG and FC. The availability of an EST collection representing single-nucleotide polymorphism variation among alleles in tetraploid roses allows for quantitative trait locus studies in segregating generations. Such phenotyped and genotyped mapping populations will provide the key to link a large portion of the unknown petal genes with functions.

## METHODS

### Plant Material

Two tetraploid rose (*Rosa hybrida*) cultivars were used throughout the experiments: Fragrant Cloud (FC) and Golden Gate (GG). Plants were grown in pots filled with a peat:volcanic gravel mixture (1:1, v/v) in the greenhouse under controlled temperature (28/20°C day/night)

and a natural photoperiod. Flower development was divided into six stages (Figure 1). At stage 1, the flower bud is closed and petals are green. At stage 2, petals start to emerge from the sepal, and their color changes to red (anthocyanin in FC) or yellow (carotenoids in GG). Stages 3 and 4 are characterized by rapid petal elongation and further accumulation of pigments. At stage 5, petals unroll, reaching full opening at stage 6.

#### RNA Isolation, cDNA Library Construction, Clone Excision, and Nucleotide Sequencing

Total RNA was isolated from rose petals or leaves of FC and GG as described previously (Manning, 1991). Poly(A)<sup>+</sup> RNA was isolated from total RNA using an Oligotex kit according to the manufacturer's instructions (Qiagen, Chatsworth, CA). cDNA synthesis from 5 µg of poly(A)<sup>+</sup> RNA, isolated from petals of FC and GG flowers at stage 4, and library construction (directional cloning) in λZAP Express vector were performed according to the manufacturer's instructions (Stratagene, La Jolla, CA). cDNA libraries of  $2.1 \times 10^6$  plaque forming units for FC and  $2.5 \times 10^6$  plaque forming units for GG were obtained. To estimate the average insert size, pBK-CMV phagemid vectors containing the inserts were excised *in vivo* from 50 individual ZAP vectors according to the manufacturer's instructions (Stratagene). Plasmid DNA was isolated with a mini-prep kit (Qiagen) and used as a template for PCR amplification of the inserts using T3 and T7 primers. The average size of the amplified product in both libraries, as analyzed by agarose gel electrophoresis, was ~1.3 kb. To perform sequence analysis of individual ESTs, ~3 × 10<sup>3</sup> randomly chosen plaques (2000 FC and 1000 GG) were isolated from the library after amplification and excised individually as described above. High-quality plasmid DNA was extracted and inserts were sequenced at the Weizmann Institute and at Lark Technologies (Saffron Walden, UK) using BioDye terminator cycle sequencing reactions and T3 primer (5' end). For selected clones, the complete sequence of both strands was obtained using T3, T7, and sequence-specific primers.

#### Sequence Analysis

Sequences were analyzed from trace data (ABI chromatograms) using PHRED software (<http://bozeman.mbt.washington.edu/phrap.docs/phred.html>) for base calling and assigning quality values. The output sequences were edited both manually and using TIGR's LUCY software (<http://www.tigr.org/softlab/>) to remove sequences originating from vectors and poor-quality 3' sequences. Contigs were assembled from the edited ESTs using PHRAP software (<http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>). All petal ESTs were compared with GenBank protein sequences using the BLASTX algorithm (NR, the nonredundant proteins database). The TBLASTX algorithm was used to compare all six-frame translations of petal ESTs with all six-frame translations of plant ESTs (EST\_pln database). TBLASTX also was used to compare petal ESTs with the Arabidopsis genomic sequences using the Arabidopsis "tilingpath" local database developed by the Solanaceae Genomic Network team at Cornell University ([http://www.sgn.cornell.edu/about/at\\_path.html](http://www.sgn.cornell.edu/about/at_path.html)).

#### Preparation and Arraying of the Target cDNAs

ESTs from FC were selected after annotation, and the corresponding cDNA clones were amplified by PCR (M.J. Research, Inc., Water-

town, MA) using T3 and T7 primers. PCR (100 µL total volume) was prepared in a 96-well plate (Abgene, Surrey, UK), and amplified products were purified with the 96-well PCR purification kit (Elge BioSystems, Gaithersburg, MD). After purification, the samples were dried, resuspended in 15 µL of 50% DMSO (100 to 200 ng DNA/µL), and transferred to a 384-format plate. Amplified cDNAs were applied to poly-L-Lys-coated glass slides (British Drug House, Dorset, UK) using an arraying robot with 16-pin print heads (BioRobotics, Ltd., Cambridge, UK). Each cDNA clone was arrayed three times at random positions. For controls, yeast cDNA, *clb3*, and DMSO without DNA were arrayed. The slides were air-dried and then baked for 2 h at 80°C. cDNAs were UV cross-linked, and slides were blocked with 6 g of succinic anhydride dissolved in 325 mL of 1-methyl-2-pyrrolidinone and 15 mL of 1 M sodium borate, pH 8.0, for 15 min. Printed DNA was denatured by washing the slides in DDW at 95°C for 2 min and then in 95% ethanol for 1 min. Slides were dried by centrifugation.

#### Preparation of Fluorescent Probes

Total RNA was extracted from FC and GG petals at stages 1 and 4; mRNA was isolated using an Oligotex mRNA purification kit (Qiagen), and reverse transcription was performed. cDNA was labeled indirectly with succinimidyl ester Cy3/Cy5 as follows. Petal mRNA (2 µg) and 2 ng of *in vitro*-synthesized yeast RNA for *clb3* (control) were combined with 4 µg of Oligo(dT) 12- to 18-mer and 4 µg of random hexamers (Amersham Pharmacia Biotech, Buckinghamshire, UK) in a final volume of 15 µL and heated to 70°C for 10 min to denature RNA tertiary structures, then the mixture was transferred to 45°C to anneal the primers. A preheated (50°C) 30-µL mixture containing 1 × RT buffer (Gibco BRL, Invitrogen, Ltd., Paisley, UK), 50 mM MgCl<sub>2</sub>, 0.1 M DTT (Gibco BRL, Invitrogen, Ltd.), 20 units of RNasein (Promega, Madison, WI), and 10 mM deoxynucleotide triphosphate mixture (2.5 mM each of dATP, dCTP, and dGTP, 1.5 mM dTTP, and 1 mM aminoallyl dUTP) was added to the RNA, and the reaction mixture was incubated for 1 h at 45°C with 400 units of Superscript II (Gibco BRL, Invitrogen, Ltd). An additional 400 units of the enzyme was applied to the reaction mixture, which was incubated for another 1 h. After incubation, RNA was hydrolyzed by adding 10 µL of 1 M NaOH and incubating at 65°C for 20 min. The reaction mixture was neutralized by titration with 1 M Tris-HCl, pH 7.5, and DNA was precipitated with ethanol.

Aminoallyl-labeled cDNA was resuspended in 10 µL of 0.1 M NaHCO<sub>3</sub>/Na<sub>2</sub>CO<sub>3</sub> buffer, pH 9.3, and aliquots (4.5 µL in DMSO) of commercial succinimidyl ester Cy3/Cy5 dyes (Amersham Pharmacia Biotech) were added according to the manufacturer's instructions. After 1 h of incubation in the dark at room temperature, Cy3/Cy5-labeled DNA was purified using a PCR purification kit (Qiagen). The ratios of light absorption at A<sub>550</sub>/A<sub>260</sub> and A<sub>650</sub>/A<sub>260</sub> were measured using a V-570 spectrophotometer (Jasco Corp., Tokyo, Japan) to evaluate the incorporation of Cy3 and Cy5 into the DNA.

#### Microarray Hybridization and Scanning

Slides were incubated for 45 min at 42°C in preheated prehybridization buffer (5 × SSC [1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate], 0.1% SDS, and 1% BSA), washed in water and isopropanol, and dried by centrifugation. The probes (Cy3- and Cy5-labeled cDNA) were mixed, dissolved in 12 µL of 2 × hybridization buffer (final concentration of 25% formamide, 5% SSC, and 0.1% SDS)

containing 15  $\mu\text{g}$  of poly(A) DNA (Amersham Pharmacia Biotech) and 5  $\mu\text{g}$  of yeast tRNA (Gibco BRL), denatured at 95°C for 3 min, and cooled on ice for 30 s. The probe (22  $\mu\text{L}$ ) was applied directly to the slides (under a cover slip), which then were incubated in a hybridization chamber (TeleChem International, Sunnyvale, CA) at 42°C for 20 h. After hybridization, the slides were washed with  $2 \times \text{SSC}$  and 0.1% SDS for 5 min at 42°C, with  $0.1 \times \text{SSC}$  and 0.1% SDS for 10 min at room temperature, and then with  $0.1 \times \text{SSC}$  for 1 min, four times at room temperature, and finally dried quickly by centrifugation. Hybridized microarray slides were scanned immediately for fluorescence emission using ScanArray Lite (Packard BioScience, PerkinElmer Life Sciences, Inc., Boston, MA). Laser and photomultiplier voltages were adjusted manually to minimize the number of spots returning saturated values. The scanner output images were quantified using QuantArray analysis software (Packard BioScience, PerkinElmer Life Sciences, Inc.).

### Microarray Data Analysis

Spot intensities were quantified using Imagene 4.2 image-analysis software (BioDiscovery, Inc., Marina del Rey, CA), and channel ratios were determined by the median-of-ratio method. Signals were normalized by dividing the intensity of each spot by the average intensity of all control spots on the same slide ("slide-wise normalization"; Schuchhardt et al., 2000). A quality-control procedure was conducted before data from the six replicates of the two independent arrays were averaged. All low-quality spots and/or spots showing large differences between the ratio mean and the ratio median were removed from the final analysis. Finally, only spots that exhibited signals higher than those of the array backgrounds for both channels and whose signals were twofold higher than the background for at least one channel were analyzed further.

### RNA Gel Blot Analysis

Total RNA (10  $\mu\text{g}$ ) was fractionated on a 1% agarose gel containing formaldehyde and blotted onto Hybond N<sup>+</sup> membranes (Amersham Pharmacia Biotech). The blots were hybridized in 0.263 M Na<sub>2</sub>HPO<sub>4</sub>, 7% SDS, 1 mM EDTA, and 1% BSA at 60°C with <sup>32</sup>P-labeled specific cDNA probes (Rediprime; Amersham Pharmacia Biotech). The membranes were washed twice in  $2 \times \text{SSC}$  and 0.1% SDS at 60°C for 20 min each and exposed to x-ray film (Fuji, Tokyo, Japan) with two intensifying screens at −70°C.

### Collection of Headspace Volatiles and Gas Chromatography–Mass Spectrometry Analysis

The headspace volatiles of intact rose flowers were trapped, eluted, and concentrated as described previously (Lavid et al., 2002). The volatile compounds collected from the headspace were analyzed on an HP-GCD apparatus equipped with an Rtx-5SIL mass spectrometer (Restek, Waldbronn, Germany; i.d. of 0.25  $\mu\text{m}$ , 30 m  $\times$  0.25 mm) fused silica capillary column. Helium (0.8 mL/min) was used as the carrier gas. The injector temperature was 250°C, set for splitless injection. The oven was set to 50°C for 1 min, and then the temperature was increased to 260°C at a rate of 5°C/min. The transfer line temperature was 280°C. Mass range was recorded from 42 to 450 mass-to-charge ratio, with an electron energy of 70 eV. The main components were identified by comparing mass spectra and retention time

data with those of authentic available samples supplemented with NIST98 and Quadlib 1607 libraries. Biosynthetic germacrene D was identified by comparison with the Kovac index of authentic germacrene D and by analysis of its mass spectrum using a computerized NIST98 library. These properties were compared with those of authentic germacrene D present in the commercial essential oil of *Cananga odorata* (Roth, Karlsruhe, Germany). For quantitative analyses, 10  $\mu\text{g/mL}$  ethyl myristate was used as an internal standard (Shalit et al., 2001; Lavid et al., 2002).

### Expression of cDNA in *Escherichia coli*

The putative sesquiterpene synthase (EST FC0592) was cloned into TOPO expression vector (Invitrogen, Paisley, UK) and transformed into *E. coli* BL21(DE3) pLysS. Bacterial lysate, prepared after isopropylthio- $\beta$ -galactoside induction (Lavid et al., 2002), was used to assay the enzymatic activity of the recombinant gene product.

### Enzymatic Assays

Farnesyl diphosphate (1  $\mu\text{M}$ ; Sigma, St. Louis, MO), 10 mM MgCl<sub>2</sub>, and 400  $\mu\text{L}$  of bacterial lysate were mixed in a total volume of 2 mL of reaction buffer, overlaid with 2 mL of hexane, and incubated overnight at 30°C. Each tube was shaken and extracted sequentially (three to five times) with 2 mL of hexane. The hexane layers containing the in vitro–formed sesquiterpenes were pooled, passed through a small Pasteur pipette filled with silica gel, dried with sodium sulfate, and concentrated to a final volume of 400  $\mu\text{L}$ . A 1- $\mu\text{L}$  aliquot was injected into the gas chromatography–mass spectrometry system for the identification of volatiles formed (Shalit et al., 2001).

Upon request, all novel materials described in this article will be made available in a timely manner for noncommercial research purposes. No restrictions or conditions will be placed on the use of any materials described in this article that would limit their use for noncommercial research purposes.

### Accession Numbers

The EST sequences were submitted to the GenBank (dbEST) database (accession numbers BQ103855 to BQ106728). Accession numbers for other sequences mentioned are cotton (+)- $\delta$ -cadinene synthase (CAA77191) and tomato germacrene C synthase (AAC39432).

### ACKNOWLEDGMENTS

We thank Robert Fluhr for helping us with the microarray analyses, Eitan Rubin and Ron Ofir for assisting us with the statistical analyses of the microarray results, Uzi Ravid and Olga Larkov for their help in the identification of germacrene D, and Dvorah Weisman for assisting us with computerized analyses of data. E.P. was the recipient of a U.S. Department of Agriculture/U.S.–Israel Binational Agricultural Research and Development Fund scholarship and a Forchheimer Visiting Professor Fellowship from the Hebrew University of Jerusalem. This research was supported by the Israeli Ministry of Science, Culture, and Sport (Grant 1410).

Received June 10, 2002; accepted July 28, 2002.

## REFERENCES

- Aharoni, A., et al. (2000). Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell* **12**, 613–616.
- Back, K., and Chappell, J. (1995). Cloning and bacterial expression of a sesquiterpene cyclase from *Hyoscyamus muticus* and its molecular comparison to related terpene cyclases. *J. Biol. Chem.* **270**, 7375–7382.
- Barletta, A. (1995). Scent makes a comeback. *Floraculture* **5**, 23–25.
- Ben-Nissan, G., and Weiss, D. (1996). The petunia homologue of tomato *gast1*: Transcript accumulation coincides with gibberellin-induced corolla cell elongation. *Plant Mol. Biol.* **32**, 1067–1074.
- Bohlman, J., Gershenzon, J., and Aubourg, S. (2000). Biochemical, molecular genetic and evolutionary aspects of defense-related terpenoid metabolism in conifers. In *Evolution of Metabolic Pathways, Recent Advances in Phytochemistry*, Vol. 34, J.T. Romeo, R. Ibrahim, L. Varin, and V. De Luca, eds (New York: Pergamon Press), pp. 109–150.
- Channeliere, S., Riviere, S., Scalliet, G., Jullien, F., Szecsi, J., Dolle, C., Vergne, P., Dumas, C., Bendahmane, M., Hugueney, P., and Cock, J.M. (2002). Analysis of gene expression in rose petals using expressed sequence tags. *FEBS Lett.* **515**, 35–38.
- Chen, X.Y., Chen, Y., Heinsteins, P., and Davisson, V.J. (1995). Cloning, expression, and characterization of (+)-delta-cadinene synthase: A catalyst for cotton phytoalexin biosynthesis. *Arch. Biochem. Biophys.* **324**, 255–266.
- Colby, S.M., Crock, J., Dowdle-Rizzo, B., Lemaux, P.G., and Croteau, R. (1998). Germacrene C synthase from *Lycopersicon esculentum* cv. VFNT Cherry tomato: cDNA isolation, characterization, and bacterial expression of the multiple product sesquiterpene cyclase. *Proc. Natl. Acad. Sci. USA* **95**, 2216–2221.
- Croteau, R., and Karp, F. (1991). Origin of natural odorants. In *Perfume: Art, Science and Technology*, P. Muller and D. Lamparsky, eds (New York: Elsevier Applied Sciences), pp. 101–126.
- Dudareva, N. (2002). Molecular control of floral fragrance. In *Breeding for Ornamentals: Classical and Molecular Approaches*, A. Vainstein, ed (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 295–309.
- Dudareva, N., Cseke, L., Blanc, V.M., and Pichersky, E. (1996). Evolution of floral scent in *Clarkia*: Novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *Plant Cell* **8**, 1137–1148.
- Dudareva, N., D'Auria, J.C., Nam, K.H., Raguso, R.A., and Pichersky, E. (1998). Acetyl-CoA:benzylalcohol acetyltransferase: An enzyme involved in floral scent production in *Clarkia breweri*. *Plant J.* **14**, 297–304.
- Dudareva, N., Murfitt, L.M., Mann, C.J., Gorenstein, N., Kolosova, N., Kish, C.M., Bonham, C., and Wood, K. (2000). Developmental regulation of methyl benzoate biosynthesis and emission in snapdragon flowers. *Plant Cell* **12**, 949–961.
- Dudareva, N., and Pichersky, E. (2000). Biochemical and molecular genetic aspects of floral scents. *Plant Physiol.* **122**, 627–633.
- Facchini, P.J., and Chappell, J. (1992). Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. *Proc. Natl. Acad. Sci. USA* **89**, 11088–11092.
- Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R.N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **18**, 1157–1161.
- Flament, I., Debonneville, C., and Furrer, A. (1993). Volatile constituents of roses: Characterization of cultivars based on the headspace analysis of living flower emissions. In *Bioactive Volatile Compounds from Plants*, R. Teranishi, R.G. Buttery, and H. Sugisawa, eds (Washington, DC: American Chemical Society), pp. 269–281.
- Gang, D., Wang, J., Dudareva, N., Hee Nam, K., Simon, J.E., Lewinsohn, E., and Pichersky, E. (2001). An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiol.* **125**, 539–555.
- Gudin, S. (2000). Rose: Genetics and breeding. *Plant Breeding* **17**, 159–189.
- Guterman, I., et al. (2002). An integrated genomic approach to discovering fragrance-related genes in rose petals. *Flowering Newsl.* **32**, 31–37.
- Knudsen, J.T., Tollesten, L., and Bergstrom, G.L. (1993). Floral scent: A checklist of volatile compounds isolated by head-space techniques. *Phytochemistry* **33**, 252–280.
- Kolosova, N., Sherman, D., Karlson, D., and Dudareva, N. (2001). Cellular and subcellular localization of S-adenosyl-L-methionine:benzoic acid carboxyl methyltransferase, the enzyme responsible for biosynthesis of the volatile ester methylbenzoate in snapdragon flowers. *Plant Physiol.* **126**, 956–964.
- Lange, B.M., Wildung, M.R., Stauber, E.J., Sanchez, C., Pouchnik, D., and Croteau, R. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proc. Natl. Acad. Sci. USA* **97**, 2934–2939.
- Lavid, N., et al. (2002). O-Methyltransferases involved in the biosynthesis of volatile phenolic derivatives in rose petals. *Plant Physiol.* **129**, 1899–1907.
- Ma, L., Li, J., Qu, L., Hager, J., Chen, Z., Zhao, H., and Deng, X.W. (2001). Light control of Arabidopsis development entails coordinated regulation of genome expression and cellular pathways. *Plant Cell* **13**, 2589–2607.
- Manning, K. (1991). Isolation of nucleic acids from plants by differential solvent precipitation. *Anal. Biochem.* **195**, 45–50.
- Martin, C., and Gerats, T. (1993). Control of pigment biosynthesis genes during petal development. *Plant Cell* **5**, 1253–1264.
- Pichersky, E., and Gang, D.R. (2000). Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. *Trends Plant Sci.* **5**, 439–445.
- Sablowski, R.W.M., Moyano, E., Culianezmacia, F., Schuch, W., Martin, C., and Bevan, M.A. (1994). Flower-specific myb protein activates transcription of phenylpropanoid biosynthetic genes. *EMBO J.* **13**, 128–137.
- Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., and Herzel, H. (2000). Normalization strategies for cDNA microarrays. *Nucleic Acid Res.* **28**, e47.
- Schünmann, P.H.D., Smith, R.C., Lång, V., Matthews, P.R., and Chandler, P.M. (1997). Expression of xet-related genes and its correlation to elongation in leaves of barley (*Hordeum vulgare* L.). *Plant Cell Environ.* **20**, 1439–1450.
- Shalit, M., Katzir, N., Tadmor, Y., Larkov, O., Burger, Y., Schalechet, F., Lastochkin, E., Ravid, U., Amar, O., Edelstein, M., and Lewinsohn, E. (2001). Acetyl CoA: Alcohol acetyl transferase activity and aroma formation in ripening melon fruits. *J. Agric. Food Chem.* **49**, 794–799.
- Shi, L., Gast, R.T., Golparaj, M., and Olszewski, N.E. (1992).

- Characterization of a shoot-specific, GA<sub>3</sub>- and ABA-regulated gene from tomato. *Plant J.* **2**, 623–632.
- Sterky, F., et al.** (1998). Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. *Proc. Natl. Acad. Sci. USA* **95**, 13330–13335.
- Uimari, A., and Strommer, J.** (1997). Myb26: A MYB-like protein of pea flowers with affinity for promoters of phenylpropanoid genes. *Plant J.* **12**, 1273–1284.
- Vainstein, A., Lewinsohn, E., Pichersky, E., and Weiss, D.** (2001). Floral fragrance: New inroads into an old commodity. *Plant Physiol.* **127**, 1383–1389.
- Vishnevetsky, M., Ovadis, M., and Vainstein, A.** (1999). Carotenoid sequestration in plants: The role of carotenoid-associated proteins. *Trends Plant Sci.* **4**, 232–235.
- Weiss, D.** (2000). Regulation of flower pigmentation and growth: Multiple signaling pathways control anthocyanin synthesis in expanding petals. *Physiol. Plant.* **10**, 152–157.
- Weiss, E.A.** (1997). Rosaceae. In *Essential Oil Crops*. E.A. Weiss, ed (Oxon, UK: CAB International Wallingford), pp. 393–416.
- Zamir, D.** (2001). Improving plant breeding with exotic genetic libraries. *Nat. Rev. Genet.* **2**, 983–989.
- Zuker, A., Tzfira, T., Ben-Meir, H., Ovadis, M., Shklarman, E., Itzhaki, H., Forkmann, G., Martens, S., Neta-Sharir, I., Weiss, D., and Vainstein, A.** (2002). Suppression of anthocyanin synthesis by antisense *fht* enhances flower fragrance. *Mol. Breeding* **9**, 33–41.
- Zuker, A., Tzfira, T., and Vainstein, A.** (1998). Cut-flower improvement using genetic engineering. *Biotechnol. Adv.* **16**, 33–79.